Calibrating Arbitrage-Free Stochastic Volatility Models by Relative Entropy Method [1]

R. Carmona [2] and Lin Xu[2]
Statistics & Operations Research Program, C.E.O.R.
Princeton University
Princeton, NJ 08544, USA

July 1, 1997

[1]AMS 1991 subject classification: 60H25
Key Words and Phrases: Option pricing, implied volatility, volatility smile and skew, stochastic volatility model, relative entropy distance, stochastic control,
Running Title: Relative Entropy Calibration
[2]Partially supported by ONR N00014-91-1010

Abstract

We develop a new framework to calibrate stochastic volatility option pricing models to an arbitrary prescribed set of prices of liquidly traded options. Our approach produces an arbitrage-free stochastic volatility diffusion process that minimizes the distance to a prior diffusion model. We use the notion of relative entropy (also known under the name of Kullback-Leibler distance) to quantify the distance between the two diffusions. The problem is formulated as a stochastic control problem.
We also show that, in a very natural limiting regime, it results in a calibrating method for complete models. Implementation issues are discussed in details for calibrating both the stochastic volatility and the complete models.

1 Introduction

One of the most successful theoretical achievements in practical finance is the Black-Scholes formula. It allows one to compute the prices of European options at different maturities and strikes provided that the volatility of the underlying stock is known. Conversely, this very formula can be used to calculate the volatility of the underlying stock from the market quote of an European option. The volatility computed in this way is called the implied volatility. If the market had the good taste to follow the Black-Scholes theory, the implied volatilities obtained from different options with different maturities and strikes should be the same. However, real life *does not* support this implication of the Black-Scholes theory. For example, it is well known that out of the money $S\&P$ 500 index put options are traded at a higher volatility than in the money puts. In many option markets, say currency option markets for example, implied volatilities exhibit a "smile" and a " skew" in both maturity and strike. To model situations where the implied volatility depends both on the time to maturity and the strike in such a way, many researchers have proposed new types of arbitrage-free diffusion models for the underlying asset. One such type is a complete diffusion models in which the spot volatility is a deterministic function of time and the underlying asset price level. Another type is a stochastic volatility model in which the spot volatility itself is subject to independent random shocks. The object then is to estimate what the volatility structure should be from the available observed liquid option prices so that one can systematically price other derivatives and in particular OTC options.

Here we provide a simple quantitative approach for constructing such an arbitrage-free diffusion with stochastic volatility. As a result of a natural limit procedure, the standard procedure used in the framework of complete models is also recovered.

The basic idea is as follows. We use the following diffusion model to capture the variability of the underlying asset price and its volatility:

$$\left\{ \begin{array}{lcl} d\,S_t & = & S_t[\alpha dt + \sqrt{g(Y_t)}dB_t] \\ dY_t & = & \mu(t, \log S_t, Y_t)\; dt + dZ_t, \end{array} \right. \tag{1}$$

where S_t denotes the price at time t of the underlying asset, B_t and Z_t are two independent Brownian motions and where g is a nonnegative function bounded away from 0 and ∞. Its role is to define a reasonable volatility from the semi-martingale Y_t which could be negative at times. In most applications we choose g to be of the form:

$$g(y) = a_1 + \frac{a_2 - a_1}{\pi}(\arctan(y) + \frac{\pi}{2}) \tag{2}$$

for a pair of constants $0 < a_1 < a_2 < \infty$ which play the roles of rough lower and upper bounds for the instantaneous volatility of the underlying asset.) The constant α is the rate of return on the underlying asset in the risk-neutral world (for example it is equal to the risk-free interest rate r for domestic assets and to the domestic risk-free rate after subtraction of the relevant foreign interest rate for currencies.) The assumptions on the drift term μ will be made explicit later in the paper. At this stage, we should merely assume that there is existence and uniqueness for the diffusion process $\{(X_t, Y_t);\ t \geq 0\}$ given by the model (1). We shall use the notation $\mathbb{P}^\mu$ for the distribution of the diffusion given by (1) and we shall denote by $\mathbb{E}^\mu$ the corresponding expectation.

As we shall see later, the above form of the model is very rich. In particular, a specific approximation procedure makes it possible to include complete models as well.

Our approach is Bayesian in spirit in the sense that we use a prior distribution for the evolution of the underlying asset in the risk-neutral world. That is we start with a version:

$$\begin{cases} d\,S_t^0 &= S_t^0[\alpha dt + \sqrt{g(Y_t)} dB_t] \\ dY_t^0 &= \mu_0(t, \log S_t^0, Y_t^0)\; dt + dZ_t. \end{cases}$$

of the model (1) in which all the coefficients are known and we use all the available information to derive a realistic form of the model (1) by demanding that it prices correctly all existing liquid European options while keeping the corresponding diffusion as close to the prior as possible. The notion of closeness is to be understood in the sense of a distance (or at least pseudo - distance) between diffusion processes. We use the relative entropy also known as the Kullback-Leibler information distance (KL distance for short) to measure the distance between a candidate and the prior.

To formalize precisely the idea of correctly pricing the existing options, we introduce constraints on the volatility drift μ and we restrict our attention to the corresponding subclass of diffusions processes. We assume that we are given m expiration dates $T_1, \cdots, T_m$ and m strike prices $C_1, \cdots, C_m$ and we denote by Λ the space of functions $\mu(t, \log S, y)$ for which the diffusion process (S_t, Y_t) solution of (1) satisfies the constraints:

$$\mathbb{E}^\mu\{e^{-rT_k} f_k(\log S_{T_k})\} = C_k, \qquad k = 1, \cdots, m. \tag{3}$$

In other words, the diffusion satisfies the constraints if the m liquid options are correctly priced. For later convenience we choose to write the payoff functions f_k in terms of the logarithms of the price S instead of the prices themselves. We then look for the function μ in Λ which gives rise to a distribution $\mathbb{P}^\mu$ which minimizes the KL distance to the distribution corresponding to the prior μ_0. In other words we try to solve the constrained minimization problem:

$$\arg \inf_{\mu \in \Lambda} H(\mathbb{P}^\mu | \mathbb{P}^{\mu_0}) \tag{4}$$

where we used the notation $H(P|Q)$ for the KL pseudo-distance from P to Q. The crux of our approach is that the distance appearing in the formulation of the minimization problem (4) can be computed explicitly. Indeed, Girsanov's formula implies that:

$$H(\mathbb{P}^\mu | \mathbb{P}^{\mu_0}) = \mathbb{E}^\mu \left\{ \int_0^T [\mu(t, \log S_t, Y_t) - \mu_0(t, \log S_t, Y_t)]^2 \; ds \right\}. \tag{5}$$

This is in contrast with the case considered in [1] where the entropy distance was used only as a suggestive rationale. Indeed, the KL distance derived from the model used in [1] is identically infinite. Fortunately, as it occurs very often (for example in statistical mechanics and in large deviation theory) this entropy divergence is linear in the size of the approximation domain and the authors of [1] used this fact to get a renormalized form of the KL distance. Unfortunately, there is still an unpleasant arbitrariness in the renormalization procedure used in [1] and it is not clear what really depends upon the particular renormalization procedure. This ambiguity was one of the main motivation at the origin of the present study.

The control set Λ is chosen to incorporate in our model some of the features which the market modeler demands. In the general setting described above, the resulting stochastic control problem could be too difficult for the technology available today. In this paper, we develop the general theory first and we consider the implementation issues (down to the gory details of the numerical computations) in two specific models of importance the second one including a mean reversion component. Once the stochastic control is solved, both theoretically and numerically, we use its solution and the corresponding stochastic system (1) to price other derivatives on the same underlying asset.

We can regard, the complete model:

$$d\, S_t = S_t[\alpha dt + \sigma(t, S_t) dB_t]$$

where $h_1 \leq \sigma \leq h_2$, as a limit when $\kappa \to \infty$ of a system (1) with $\mu = \kappa[\sigma(t, S) - g(y)]$. For this very reason our results also provide a solution to the calibration problem in the framework of complete models.

Due to its important role in practical financial risk management, the subject has drawn great attention both from scholars and experts from practical finance. Since the pioneering contribution [5] of Breenden and Litzenberger in 1978, there have been many important works concerning complete models. See for example [10, 8, 6, 19, 17, 2, 1] and the references therein. We refer to [16, 20, 15, 18] and to the references therein for stochastic volatility models in the spirit of the present study. Though our work shares some same features with others, our approach is, except for [1], very different from those which we quoted. If we compare it to the approach of [1], the main difference is that we use the relative entropy in the framework of stochastic volatility models in which it makes perfectly good sense without having to rely on an ad-hoc renormalization procedure. Moreover, our approach also lends itself to the analysis of complete models. Indeed a natural limiting procedure shows that our framework contains at least in spirit complete models as well. This can be used to recover some of the numerical computations of [1] in a very natural way.

The rest of this paper is organized as follows. In the next section, we set up the stochastic control problem precisely, give some economic intuition for our choice of model and give the theoretical foundations of the algorithmic solution. We then develop in Section 3 the solution in the specific case of what we call the free model for which $\Lambda = \{\mu;\ \mu = H(t, \log S, y)\}$. In Section 4, we present the details for the case of a mean reversion model for which $\Lambda = \{\mu;\ \mu = \kappa(H(t, \log S, y) - g(y)),\ \kappa > 0,\ 0 < h_1 \leq H \leq h_2\}$. In Section 5, we derive complete models as limits of our mean reversion models of Section 4 as $\kappa \to \infty$ and we use this property to develop our algorithmic pricing procedure in this the framework. In the last section, we provide some numerical results which we compare to the numerical results of [1]

2 General Theory

Throughout more than twenty years of intensive effort following the introduction of options into financial markets, substantial evidence has been found that options are not redundant to the underlying assets and interest rate (see for example, see [4, 7, 9, 11, 13, 14].) It is now widely accepted that the option market create its own risk mostly because of the different levels of demand of the underlying asset and especially the demands coming from portfolio insurance activity. In any case, because of the popularity of options trading, the market is subject to a new dimension of risk (say liquid demand noise.) Therefore it is reasonable to use stochastic volatility model to describe the dynamics of the underlying asset and its instantaneous volatility.

We assume the underlying asset and its volatility follow the random dynamics given by the following system of stochastic differential equations:

$$\left\{ \begin{array}{rcl} d\,S_t &=& S_t[\gamma_t dt + \sqrt{g(Y_t)}dB_t] \\ dY_t &=& h(\mu_t, t, \log S_t, Y_t)\; dt + dZ_t, \end{array} \right.$$

where $B = \{B_t;\ t \geq 0\}$ and $Z = \{Z_t;\ t \geq 0\}$ are two independent Brownian motions, the function g is as defined in (2), h is a known deterministic function and the processes $\gamma = \{\gamma_t;\ t \geq 0\}$ and $\mu = \{\mu_t;\ t \geq 0\}$ are adapted to the filtration generated by the Brownian motions B and Z. However for the purpose of option pricing, we only need to consider the dynamics in the risk neutral world as given by:

$$\left\{ \begin{array}{rcl} d\,S_t &=& S_t[\alpha dt + \sqrt{g(Y_t)}dB_t] \\ dY_t &=& h(\mu_t, t, \log S_t, Y_t)\; dt + dZ_t, \end{array} \right. \tag{6}$$

where α is the rate of return on the underlying asset in the risk-neutral framework. We view the adapted process $\mu = \{\mu_t;\ t \geq 0\}$ as a control. Whenever convenient we shall use the notation $X_t = \log S_t$ for the logarithm of the price of the underlying asset. The fundamental model (1) can then be rewritten in the form:

$$\left\{ \begin{array}{rcl} dX_t &=& [\alpha - \frac{1}{2}g(Y_t)]dt + \sqrt{g(Y_t)}dB_t \\ dY_t &=& h(\mu_t, t, X_t, Y_t)\; dt + dZ_t. \end{array} \right. \tag{7}$$

Our strategy is based on the existence of m liquid European options available on the market with payoff $f_i(S_{T_i})$ at exercise time T_i. It is also based on the belief in a prior model for the diffusion (X_t, Y_t). We shall use the superscript ^{0} to emphasize the fact that these quantities refer to the prior model.

$$\left\{ \begin{array}{rcl} d\,X_t^0 &=& [\alpha - \frac{1}{2}g(Y_t^0)]dt + \sqrt{g(Y_t^0)}dB_t] \\ dY_t^0 &=& h_0(t, X_t^0, Y_t^0)\; dt + dZ_t, \end{array} \right.$$

Our terminology of *prior distribution* strongly suggests that we are heading toward a Bayesian analysis of the volatility structure of the underlying asset. Our objective is:

> to pin down for each fixed $T > 0$, a stochastic control process $\mu = \{\mu_t;\ 0 \leq t \leq T\}$ which minimizes the relative entropy distance between the distributions of the vector diffusions, i.e. the law $\mathbb{P}_{T,\mu}$ of $\{(X_t, Y_t) : 0 \leq t \leq T\}$ and the law $\mathbb{P}_T^0$ of $\{(X_t^0, Y_t^0) : 0 \leq t \leq T\}$, under the constraints that it leads to a set of correct prices for all liquid options in the market, namely such that:
>
> $$\mathbb{E}^\mu \left\{ \exp(-rT_i) f_i(X_{T_i}) \right\} = C_i \quad i = 1, \cdots, m \tag{8}$$

Using the particular form of the model (1) and the independence of the Brownian motions B and Z, it is easy to see that the probability measures $\mathbb{P}_{T,\mu}$ and $\mathbb{P}_T^0$ are absolutely continuous and that:

$$\frac{d\mathbb{P}_{T,\mu}}{d\mathbb{P}_T^0} = \exp\left[\int_0^T [(h(\mu_t, t, X_t, Y_t) - h_0(t, X_t^0, Y_t^0)]dZ_t - \frac{1}{2}\int_0^T [h(\mu_t, t, X_t, Y_t) - h_0(t, X_t, Y_t)]^2\; dt, \right]$$

and consequently:

$$H(\mathbb{P}_{T,\mu} | \mathbb{P}_T^0) = \frac{1}{2}\mathbb{E}\left\{ \int_0^T [h(\mu_s, s, X_s, Y_s) - h_0(s, X_s, Y_s)]^2\; ds \right\}.$$

Hence, finding a control μ which minimizes the relative entropy distance for the diffusions pricing correctly the liquid options is nothing but a constrained stochastic control problem. Using Lagrange multipliers $\lambda_1, \cdots, \lambda_m$ to include the constraints in the objective function our problem can indeed be reformulated as a classical stochastic control problem for a diffusion in $\mathbb{R}^2$. See for example [12] for a detailed account of this theory.

Lemma 2.1 *If we set*

$$\begin{aligned} V(t,x,y,\lambda_1,\cdots,\lambda_m) &= \sup_\mu \mathbb{E}^\mu_{t,x,y}\left\{ -\frac{1}{2}e^{rt}\int_t^T [h(\mu_s,s,X_s,Y_t) - h_0(s,X_s,Y_s)]^2\,ds \right. \\ &\qquad \left. + \sum_{t<T_j\le T} \lambda_j e^{-r(T_j-t)} f_j(\log S_{T_j}) \right\}, \end{aligned}$$

where the notation $\mathbb{E}^\mu_{t,x,y}$ is used for the distribution of the diffusion (??) conditioned at $X_t = x$ and $Y_t = y$ and if we assume that:

$$(a) \qquad |\nabla_{t,x,y} h| \le C \tag{9}$$

$$(b) \qquad |h(v,t,x,y)| \le C(1+|v|+|x|+|y|). \tag{10}$$

for some constant $C > 0$, then the value function $V(t,x,y,\lambda_1,\cdots,\lambda_m)$ is the unique solution of following partial differential equation of the Hamilton Jacobi Bellman (HJB for short) type:

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{1}{2}g(y)\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\frac{\partial^2 V}{\partial y^2} + [\alpha - \frac{1}{2}g(y)]\frac{\partial V}{\partial x} + \\ &\quad e^{rt}\Phi(e^{-rt}\frac{\partial V}{\partial y}, t, x, y) - rV = -\sum_{t<T_i\le T}\lambda_i f_i(x)\delta(t-T_i) \end{aligned} \tag{11}$$

for $0 \le t \le T$ with terminal condition:

$$V(T,x,y) = 0, \tag{12}$$

where δ is the usual δ-function at zero and where Φ is the convex conjugate defined by:

$$\Phi(p,t,x,y) = \sup_h\{hp - \frac{1}{2}[h - h_0(t,x,y)]^2\} \tag{13}$$

By construction, $\Phi(p,t,x,y)$ is a convex function of p. The following lemma sheds some light on the dependence of $V(t,x,y,\lambda_1,\cdots,\lambda_m)$ upon $\lambda_1,\cdots,\lambda_m$.

Lemma 2.2 *For each fixed t, x and y, $V(t,x,y,\lambda_1,\cdots,\lambda_m)$ is a strictly convex function of $(\lambda_1,\cdots,\lambda_m)$. Moreover if for $k = 1,\cdots,m$ we use the notation V_k for the partial derivative of V with respect to λ_k then:*

$$\begin{aligned} &\frac{\partial V_k}{\partial t} + \frac{1}{2}g(y)\frac{\partial^2 V_k}{\partial x^2} + \frac{1}{2}\frac{\partial^2 V_k}{\partial y^2} + [\alpha - \frac{1}{2}g(y)]\frac{\partial V_k}{\partial x} + \\ &\quad \frac{\partial \Phi}{\partial p}(e^{-rt}\frac{\partial V}{\partial y}, t, x, y)\frac{\partial V_k}{\partial y} - rV_k = f_k(x)\delta(t-T_k), \end{aligned} \tag{14}$$

with the terminal condition $V_k(T,x,y) = 0$.

The above equation is exactly the Black-Scholes equation in our context with:

$$h = \frac{\partial \Phi}{\partial p}\left(e^{-rt}\frac{\partial V(t,x,y)}{\partial y}, t, x, y\right).$$

As in [1] this result can be proven with standard tools of stochastic calculus in the spirit of the verification theorems (see [12].) Since our main objective here is to present and discuss algorithm procedures to price derivative, we postpone the proofs of the previous two lemmas to the Appendix at the end of the paper.

In the present setting, the calibration constraints (8) imposed by the market become:

$$\begin{array}{rcl} V_1(0,x_0,y_0) = \dfrac{\partial V}{\partial \lambda_1}(0,x_0,y_0) & = & C_1 \\ \cdots\cdots\cdots\cdots\cdots & & \cdots \\ V_m(0,x_0,y_0) = \dfrac{\partial V}{\partial \lambda_m}(0,x_0,y_0) & = & C_m \end{array}$$

2.0.1 Implementation

The theoretical facts derived above suggest the computational algorithm which we present in this section. We follow the steps of [1] where the implementation of the same idea was given in the context of complete models.

In the risk - neutral framework where we can price derivatives by expectations, the value at time $t = 0$ of a derivative with expiration T_0 and payoff function $f_0(\log S)$ when $S_0 = e^{x_0}$ and $Y_0 = y_0$ is given by:

$$e^{-rT_0}\mathbb{E}_{0,x_0,y_0}\{f_0(X_{T_0})\}. \tag{15}$$

But because of the choice of model (7 and the fact that the optimal control μ_t is such that the drift term $h(\mu_t, t, X_t, Y_t)$ is Markovian in the sense that it is a deterministic function $\tilde{h}$ of t, X_t and Y_t given by (??), the function $\varphi(t,x,y) = \mathbb{E}_{0,x,y}\{f_0(X_t)\}$ is the unique solution of the (forward) parabolic equation:

$$\frac{\partial \varphi}{\partial t} = \frac{1}{2}g(y)\frac{\partial^2 \varphi}{\partial^2 x} + \frac{1}{2}\frac{\partial^2 \varphi}{\partial^2 y} + [\alpha - \frac{1}{2}g(y)]\frac{\partial \varphi}{\partial x} + \tilde{h}(t,x,y)\frac{\partial \varphi}{\partial y} \tag{16}$$

with initial condition $\varphi(0,x,y) \equiv f_0(x)$. Consequently the various components of a computer implementation can be structured as follow:

- a PDE solver to compute simultaneously for each fixed $(\lambda_1, \cdots, \lambda_m)$ the solution

  $$\{V(t,x,y); V_k(t,x,y), k = 1, \cdots, m\}$$

  of the backward parabolic equations (11) and (14) from their final values. We give in the next sections the details of an explicit *home grown* finite difference scheme where the partial derivatives in x and y are approximated by their finite differences counterparts and the time derivative is replaced by *the backward finite difference.*

- a convex optimization solver to determine:

$$(\lambda_1^*,\cdots,\lambda_m^*) = \arg \min_{(\lambda_1,\cdots,\lambda_m)\in\mathbb{R}^m} V(0,x_0,y_0,\lambda_1,\cdots,\lambda_m) - \sum_{j=1}^m \lambda_j C_j \tag{17}$$

  when both the function to be minimized and its gradient can be computed (use the solvers alluded to in the first item with $t=0$, $x=x_0$ and $y=y_0$.) In fact as we saw in the proof of Lemma ?? the Hessian matrix could also be computed if the information on the form of the backward parabolic system (??) were to be included in the PDE solver of the first item

- use formula (??) to compute the optimal drift $\tilde{h}(t,x,y)$ from the convex conjugate function Φ and the the value function $V(t,x,y,\lambda_1^*,\cdots,\lambda_m^*)$ and solve the forward linear parabolic equation (16) to price any new derivative.

In the following we give the details of such an implementation for two specific models of importance.

3 The Free Case

This model corresponds to the choice:

$$h(v,t,x,y) = v$$

for the function h. In other words, the market has the potential to create instantaneous volatility through free choice of v based on all information up to now. Our first step is the identification of the conjugate function Φ.

$$\begin{aligned}\Phi(p,t,x,y) &= \sup_{v\in\mathcal{R}^1}\left(pv - \frac{1}{2}[v-h_0(t,x,y)]^2\right)\\ &= h_0(t,x,y)p + \frac{1}{2}p^2\end{aligned} \tag{18}$$

and since:

$$\frac{\partial\Phi}{\partial p}(p,t,x,y) = h_0(t,x,y) + p \tag{19}$$

formula (??) gives:

$$\mu_t = h_0(t,x,y) + e^{-rt}\frac{\partial V}{\partial y}(t,x,y). \tag{20}$$

The **HJB** partial differential equations (11) and (14) become completely explicit if we plug in the expressions (18) and (19) and we can describe the numerical algorithm to solve them in full detail. We use a finite difference scheme to solve the system. For this, we consider the time/space grid:

$$\{(i\Delta_N t, x_0 + j_1\delta, y_0 + j_2\delta); \Delta_N t = \frac{t}{N},\ \delta>0,\ j_1,j_2 = 0,\pm1,\pm2,\cdots\}, \tag{21}$$

where $x_0 = \log S_0$, y_0 is the same as in the prior and δ and N are to be selected later. For the sake of simplicity we use from now on the notation $(i,j_1,j_2)_P$ for the generic point $(i\Delta_N t, x_0+j_1\delta, y_0+j_2\delta)$ of our space - time grid.

The numerical solution of the partial differential equation (11) is the finite grid approximation obtained by replacing:

- $\frac{\partial V}{\partial x}(i,j_1,j_2)_P$ by $D_{x,i,j_1,j_2}V = \frac{V(i,j_1+1,j_2)_P - V(i,j_1-1,j_2)_P}{2\delta}$;
- $\frac{\partial V}{\partial x}(i,j_1,j_2)_P$ by $D_{y,i,j_1,j_2}V = \frac{V(i,j_1,j_2+1)_P - V(i,j_1,j_2-1)_P}{2\delta}$;
- $\frac{\partial V}{\partial t}(i,j_1,j_2)_P$ by $D_{t,i,j_1,j_2}V = \frac{V(i-1,j_1,j_2)_P - V(i,j_1,j_2-)_P}{-\Delta_N t}$;
- $\frac{\partial^2 V}{\partial x^2}(i,j_1,j_2)_P$ by $D_{xx,i,j_1,j_2}V = \frac{V(i,j_1+1,j_2)_P - 2V(i,j_1,j_2)_P V(i,j_1-1,j_2)_P}{\delta^2}$.
- $\frac{\partial^2 V}{\partial x^2}(i,j_1,j_2)_P$ by $D_{yy,i,j_1,j_2}V = \frac{V(i,j_1,j_2+1)_P - 2V(i,j_1,j_2-1)_P V(i,j_1,j_2-1)_P}{\delta^2}$.

For stability reasons we require that:

$$(3+a_2)\Delta_N t \leq \delta^2. \tag{22}$$

This leads to the simple iterative scheme to compute: $V(0,x_0,y_0)$:

$$\begin{aligned} V(i-1,j_1,j_2) &= (1-r\Delta_N t)V(i,j_1,j_2)_P + \Delta_N t \Bigg[\frac{1}{2}g(j_2\delta)D_{xx,i,j_1,j_2}V + \frac{1}{2}D_{yy,i,j_1,j_2}V \\ &\quad +[\alpha - \frac{1}{2}g(j_2\delta)]D_{x,i,j_1,j_2}V + h_0(i,j_1,j_2)_P D_{y,i,j_1,j_2}V + \frac{1}{2}e^{-ri\Delta_N t}(D_{y,i,j_1,j_2}V)^2 - rV \\ &\quad + \sum_{(i-1)\Delta_N t < T_i \leq j\Delta_N t} \lambda_j f_j(x)\delta(t-T_j)\Bigg] \end{aligned} \tag{23}$$

Using similar notation for the evaluation of the partial derivatives V_k and their partial derivatives with respect to t, x and y on the grid, we are led to the solution of the system of (backward) difference equations:

$$\begin{aligned} V_k(i-1,j_1,j_2) &= (1-r\Delta_N t)V_k(i,j_1,j_2)_P + \Delta_N t \Bigg[\frac{1}{2}g(j_2\delta)D_{xx,i,j_1,j_2}V_k + \frac{1}{2}D_{yy,i,j_1,j_2}V_k \\ &\quad +[\alpha - \frac{1}{2}g(j_2\delta)]D_{x,i,j_1,j_2}V_k + h_0(i,j_1,j_2)_P + e^{-ri\Delta_N t}D_{y,i,j_1,j_2}V - rV + f_k(x)\delta(i\Delta_N t - T_k)\Bigg] \end{aligned} \tag{24}$$

Once the optimal values $\{\lambda_1^*,\cdots,\lambda_m^*\}$ of the Lagrange multipliers are found one can use the calibrated model with the optimal drift:

$$\tilde{h}(i,j_1,j_2)_P = H_0(i,j_1,j_2)_P + D_{y,i,j_1,j_2}V. \tag{25}$$

to price new derivatives with the forward parabolic partial differential equation (16). But the advantage of having the optimal drift goes beyond these simple pricing procedures. Indeed this optimal drift can be used to price more complex derivatives (barrier options are the first ones to come to mind) my Monte Carlo simulation of the actual sample paths of the solution of the original system.

4 Mean Reversion Case

Quite often, market modelers require that the volatility of the underlying asset has some sort of mean reversion property. This can be accomplished in the present setting by choosing the drift function h in the following way:

$$h(p,t,x,y) = \kappa[p - g(y)], \tag{26}$$

where $a_1 < c_1 \le p \le c_2 < a_2$ and κ is a positive constant to be chosen later. In this section we assume that the prior is given by:

$$\begin{cases} d\,S_t^0 = & S_t^0[\alpha dt + \sqrt{g(Y_t^0)}dB_t] \\ dY_t^0 = & \kappa[A_0(t, \log S_t^0, Y_t^0) - g(Y_y^0)]\ dt + dZ_t, \end{cases}$$

with $a_1 < c_1 \le A_0 \le c_2 < a_2$ and we suppose that the candidates for the dynamics of the underlying asset are restricted to the solutions of the system:

$$\begin{cases} d\,S_t = & S_t[\alpha dt + \sqrt{g(Y_t)}dB_t] \\ dY_t = & \kappa[\mu_t - g(Y_t)]\ dt + dZ_t, \end{cases}$$

with $c_1 \le \mu_t \le c_2$. We look for $\{\mu_t : 0 \le t \le T\}$ minimizing:

$$\frac{1}{2}\kappa^2 \int_0^T [\mu_t - A_0(t, \log S_t, Y_t)]^2\ dt, \tag{27}$$

under constraints that it prices all the available liquid options correctly, namely

$$\mathbb{E}^\mu \left\{ e^{-rT_i} f_i(\log S_{T_i}) \right\} = C_i \qquad\qquad i = 1, \cdots, m \tag{28}$$

This optimization problem can be solved with a numerical algorithm very similar to the one used in the previous section. The only differences arise because of the special form of the convex conjugate function Φ. We compute its new value as follows.

$$\begin{aligned} \Phi(p,t,x,y) &= \sup_{c_1 \le \mu \le c_2} [\kappa(\mu - g(y))p - \frac{1}{2}\kappa^2(\mu - A_0(t,x,y))^2] \\ &= \kappa[A_0(t,x,y) - g(y)]p + \sup_{\kappa(c_1 - A_0) \le v \le \kappa(c_2 - A_0)} [\kappa v p - \frac{1}{2}v^2] \\ &= \kappa[A_0(t,x,y) - g(y)]p + \Psi_{\kappa(c_1-A_0),\kappa(c_2-A_0)}(p), \end{aligned}$$

where

$$\Psi_{d_1,d_2}(p) = \begin{cases} \frac{1}{2}p^2 & \text{if } d_1 \le p \le d_2\ , \\ d_2 p - \frac{1}{2}(d_2)^2 & \text{if } p > d_2\ , \\ d_1 p - \frac{1}{2}(d_2)^1 & \text{if } p < d_1 \end{cases}$$

Moreover the optimal μ is given by:

$$\mu = \begin{cases} A_0 + p/\kappa & \text{if } d_1 \le p \le d_2\ , \\ c_2 & \text{if } p > d_2\ , \\ c_1 & \text{if } p < d_1 \end{cases}$$

where $d_1 = \kappa(c_1 - A_0)$ and $d_2 = \kappa(c_2 - A_0)$.

5 Complete Models as Limiting Cases

In the previous section, we argue that any complete model can be reasonably regarded as a stochastic volatility model of the type considered in this paper. But independently of the fact that sometime one still would like to select calibrating candidates within complete model category. In this section, we show that any complete model appears as the limit as $\kappa \to \infty$ of a stochastic system of the form (6). Namely our candidate is

$$\begin{cases} d\,S_t = & S_t[\alpha dt + \sqrt{g(Y_t)}dB_t] \\ dY_t = & \kappa[\sigma_t^2 - g(Y_t)]\;dt + dZ_t, \end{cases}$$

and let us assume that our prior is of the form:

$$\begin{cases} d\,S_t^0 & = S_t^0[\alpha dt + \sqrt{g(Y_t^0)}dB_t] \\ dY_t^0 & = \kappa[\sigma_0^2(t, \log S_t^0) - g(Y_y^0)]\;dt + dZ_t. \end{cases}$$

For each fixed κ, minimizing the objective:

$$\frac{1}{2}\kappa^2 \int_0^T [\sigma_t^2 - \sigma_0^2(t, \log S_t)]^2\;dt$$

is obviously equivalent to minimizing:

$$\frac{1}{2}\int_0^T [\sigma_t^2 - \sigma_0^2(t, \log S_t)]^2\;dt.$$

So, when $\kappa \to \infty$, due to the mean reversion effect it is reasonable to expect that $g(Y_t) \to \sigma_t^2$. This leads to a calibration method applicable in the context of complete models:
We assume that the prior is given by the solution of the stochastic differential equation:

$$d\,S_t^0 = S_t^0[\alpha dt + \sigma_0(t, S_t^0)dB_t]. \tag{29}$$

The calibrating candidate is:

$$d\,S_t = S_t[\alpha dt + \sigma_t dB_t], \tag{30}$$

with $c_1 \leq \sigma_t^2 \leq c_2$ and we search for the control $\{\sigma_t : 0 \leq t \leq T\}$ which minimizes the functional:

$$\frac{1}{2}\int_0^T [\sigma_t^2 - \sigma_0^2(t, S_t)]^2\;dt \tag{31}$$

under the constraints of pricing correctly all the liquid options in the market. The constraints can be formulated in the form:

$$\mathbb{E}^\mu\left\{e^{-rT_i}f_i(S_{T_i})\right\} = C_i \qquad\qquad i = 1, \cdots, m. \tag{32}$$

Remark: The following is an interesting question: *if $\{S_t^\kappa : 0 \leq t \leq T\}$, that is the solution in mean reversion case of previous section, weakly converges to $\{S_t : 0 \leq t \leq T\}$ that is the solution of the above problem when $\kappa \to \infty$.*

As in Section ??, the following two lemmas provide the main tools to solve problem (??).

Lemma 5.1 *Let us set*

$$V(t,S,\lambda_1,\cdots,\lambda_m) = \sup_{\sigma} \mathbb{E}^{\sigma}_{t,S}\left\{-\frac{1}{2}e^{rt}\int_t^T(\sigma_t^2-\sigma_0^2(t,S_t))^2\,ds \right.$$
$$\left. + \sum_{t<T_i\le T}\lambda_i e^{-r(T_i-r)} f_i(\log S_{T_i})\right\}, \qquad (33)$$

where $\mathbb{E}^{\sigma}_{t,S}$ *is the law of diffusion (30) conditioned at* $S_t = S$*. Then* $V(t,S,\lambda_1,\cdots,\lambda_m)$ *is the unique solution of following* **HJB** *partial differential equation:*

$$\frac{\partial V}{\partial t} + e^{rt}\Phi\left(\frac{1}{2}e^{-rt}S^2\frac{\partial^2 V}{\partial S^2}(t,S)\right) + \alpha S\frac{\partial V}{\partial S} - rV = -\sum_{t<T_i\le T}\lambda_i f_i(S)\delta(t-T_i), \qquad (34)$$

for $0\le t\le T$ *with terminal condition:*

$$V(T+0,S) = 0. \qquad (35)$$

where δ *is the usual* δ*-function at zero, and*

$$\Phi(v,t,S) = \sup_{c_1\le u\le c_2}[uv - \frac{1}{2}[u-\sigma_0^2(t,S)]^2] \qquad (36)$$

which is convex in v*.*

Remark: Indeed a straightforward calculation gives

$$\Phi(v,t,S) = \begin{cases} \sigma_0^2(t,S)v + \frac{1}{2}v^2 & \text{if } d_1\le v\le d_2 \ , \\ c_2 v + \frac{1}{2}[c_2-\sigma_0^2(t,S)]^2 & \text{if } v > d_2 \ , \\ c_1 v + \frac{1}{2}[c_1-\sigma_0^2(t,S)]^2 & \text{if } v < d_1 \end{cases}$$

where $d_1 = c_1 - \sigma_0^2(t,S)$ and $d_1 = c_2 - \sigma_0^2(t,S)$

Lemma 5.2 $V(t,S,\lambda_1,\cdots,\lambda_m)$ *is a strictly convex function in* $(\lambda_1,\cdots,\lambda_m)$*. Moreover if we set* $V_i = \frac{\partial V}{\partial \lambda_i}$ *for* $i=1,\cdots,m$*, then*

$$\begin{aligned} &\frac{\partial V_i}{\partial t} + \frac{1}{2}\frac{\partial \Phi}{\partial v}(\frac{1}{2}e^{-rt}S^2\frac{\partial^2 V}{\partial S^2},t,S)S^2\frac{\partial^2 V_i}{\partial S^2} + \alpha S\frac{\partial V}{\partial S} \\ &= -f_i(S)\delta(t-T_i) \\ &\quad \mathit{for}\ \ 0\le t\le T, \end{aligned} \qquad (37)$$

As before we notice that this is exactly the Black-Scholes equation in the present context if we set:

$$\sigma_t^2 = \frac{\partial\Phi}{\partial v}\left(e^{-rt}S^2\frac{\partial^2 V}{\partial S^2}(t,S)\right).$$

The corresponding numerical algorithm can be described as follows:

Step 1: Picking a $(\lambda_1, \cdots, \lambda_m)$ and finding $\{V(0, S_0), V_1(0, S_0), \cdots, V_m(0, S_0)\}$. This is done by discretizing time-space variables t, S and replace derivatives by finite difference in PDE (34) and PDE (37) in the same way as in section **??** (explicit finite difference scheme). It results in an iteration algorithm which is very easy to implement.

Step 2: Based on found $\{V(0, S_0), V_1(0, S_0), \cdots, V_m(0, S_0)\}$, using a subroutine of gradient optimization to obtain a better $\{\lambda_1, \cdots, \lambda_m\}$ for the problem of minimizing

$$V(0, S, \lambda_1, \cdots, \lambda_m) - \sum_{i=1}^{m} C_1 \lambda_i.$$

Then repeat the step 1 and step 2 until satisfying $\{\lambda_1, \cdots, \lambda_m\}$ is found.

Step 3: Once the satisfying

$$\{\lambda_1, \cdots, \lambda_m\}$$

is obtained, we input it into step 1 to obtain $\{V, V_i, i = 1, \cdots, m\}$ at those time space points that are necessary to compute

$$\{V(0, S_0), V_1(0, S_0), \cdots, V_m(0, S_0)\}.$$

Along the process using

$$\sigma_t^2 = \frac{\partial \Phi}{\partial v}(e^{-rt} S^2 \frac{\partial^2 V}{\partial S^2}, t, S)$$

to obtain the optimizing σ_t^2. In fact, the finite difference algorithm used to solve the PDE (37) together with the particular value of σ_t^2 can be viewed as a tree implementation of the calibrated model. It is ready for pricing other non-liquid derivatives on the same underlying asset.

As we already mentioned a similar implementation is given in [1] and our results can be viewed as the natural extension of this philosophy to stochastic volatility models. In particular we expect that subjecting the model to the constraints of the stochastic volatility framework would result in some form of smoothing of the implied volatility structure exhibited in [1].

6 Numerical Experiments and Comments

References

[1] M. Avellaneda, C. Friedman, R. Holmes and D. Samperi (1996): Calibrating Volatility Surfaces Via Relative Entropy Minimization. (*preprint*)

[2] M. Avellaneda, A. Levy, and A. Paras (1995): Pricing and Hedging Derivative Securities in Markets with Uncertain Volatilities. *Applied Mathematical Finance,* **2**, 73-88.

[3] M. Avellaneda and A. Paras (1996): Managing the Volatility Risk of Portfolios of Derivative Securities: The Lagrangian Uncertain Volatility Model. *Applied Mathematical Finance,* **3**, 21-52.

[4] F. Black (1986): Noise. *J. of Finance,* **41**, 529-543.

[5] D.T. Breenden and R.H. Litzenberger (1978): Prices of State-Contingent Claims Implicit in Option Prices. *J. of Business,* **51**, 621-651

[6] P.W. Buchen and M. Kelly (1996): The Maximum Entropy Distribution of an Asset Inferred from Option Prices. *J. of Financial and Quantitative Analysis,* **31**, 143-159.

[7] J. Campbell, S. J. Grossman, and J. Wang (1992): Trading Volume and Serial Correlation in Stock Return. *Quarterly J. of Economics,* **108**, 905-939.

[8] E. Derman and I. Kani (1994): Riding on a Smile. *Risk Magazine,* **7**(2).

[9] B.J. DeLong, A. Shleifer, H. Summers, L. Waldmann and J. Robert (1990): Noise Trader Risk in Financial Markers. *J. of Political Economy,***98**, 703-738.

[10] B. Dupire, B. (1994): Pricing with a Smile. *Risk Magazine,***7**(1).

[11] S. Figlewski (1979): Subjective Information and Market Efficiency in a Betting Market. *Journal of Political Economics,* **87**, 75-88.

[12] W.H. Fleming and M. Soner (1992): Controlled Markov Processes and Viscosity Solutions. *Springer-Verlag,* New York.

[13] S.J. Grossman (1988): An Analysis of the Implications for Stock and Futures Price Volatility of Program Trading and Dynamic Hedging Strategies, *J. of Business,* **61**, 275-298.

[14] S.J. Grossman and Z. Zhou (1996): Equilibrium Analysis of Portfolio Insurance. *J. of Finance,* **LI(4)**, 1379-1403.

[15] S. Heston (1993): A Closed-form Solution for Options with Stochastic Volatility with Application to Bond and Currency Options. *Review of Financial Studies,* **6**, 327 -343.

[16] J. Hull and A. White (1987): The Pricing of Options on Assets with Stochastic Volatilities. *J. of Finance,* **XLII**, 281-300.

[17] J.C. Jackwerth and M. Rubinstein (1996): Recovering Probability Distributions from Contemporaneous Security Prices. *J. of Finance,* **51**, 1611-1631.

[18] E. Renault and N. Touzi (1996): Option Hedging and Implied Volatilities in a Stochastic Volatility Model. *Mathematical Finance,* **6** 279-302.

[19] M. Rubinstein (1994): Implied Binomial Trees. *J. of Finance*, **LXIX**, 771-818.

[20] E. Stein and J. Stein (1971): Stock Price Distribution with Stochastic Volatility: an Analytic Approach. *Rev. of Financial Studies*, **4**, 727-752.